Independent Peer Review Report on the

Stock Assessment Review (STAR) Panel for Yelloweye rockfish and Yellowtail rockfish

Prepared for:
The Center for Independent Experts

August 2017

Dr Kevin Stokes
Stokes.net.nz Ltd
59 Jubilee Rd
Khandallah
Wellington 6035
New Zealand
Ph: +64 (021) 222 0926
E-mail: kevin@stokes.net.nz

EXECUTIVE SUMMARY

The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

The Stock Assessment Review (STAR) Panel for yelloweye rockfish (*Sebastes ruberrimus*) and Yellowtail rockfish (*Sebastes flavidus*) was hosted by and took place at the NWFSC, Seattle, from 11th to 14th July 2017.

For all stocks, there is a problem with the lack of credible abundance indices. The AFSC Triennial and replacement NWFSC Trawl surveys both use bottom gear. Yelloweye rockfish is associated with rocky features and yellowtail is a midwater species. Neither is well sampled by bottom trawls. For yelloweye, an attempt was made to use the IPHC longline survey with additional rockfish stations. This is appropriate in principle, but in practice not wholly successful; more work is warranted. There is a similar though less consistent problem with age composition data availability. It is generally sparse and for yellowtail rockfish especially means that no age-based assessment is possible. Life history information is sufficient for yelloweye rockfish but only for the northern yellowtail stock. The attempted stock assessment for the southern yellowtail stocks borrows life history information from the north. This is not appropriate.

The yelloweye rockfish pre-STAR report is an excellent example of a comprehensive pre-STAR analysis. It was well presented and following only minor STAR input was accepted as best available information. The STAT and STAR worked together to agree on appropriate axes of uncertainty for decision tables to present to the SSC. The assessment makes well rationalized changes from the previous benchmark and update (2009/2011). Depletion at the start of 2017 is estimated as 28%, compared to 21.4% in 2011 and 16.1% in 2002). Axes of uncertainty (values of M) were chosen to characterize model specification error.

The yellowtail rockfish pre-STAR report is lacking in detail, especially with respect to data inputs. There appears to have been a problem with data availability and timeliness. The lack of detail made review difficult, but the STAT was highly responsive and the STAR was able to reach conclusions on both the northern and southern stock assessments. Lack of indices and age data, plus no stock-specific life history information makes the southern assessment untenable. More worrying is the lack of a credible index which might inform a simpler assessment approach. The northern stock assessment is again hampered by a lack of a credible abundance index but the final, recommended assessment is robust, and uncertainty is well captured in the decision tables by a range of M values.

The assessments of northern and southern Yellowtail rockfish stocks were all implemented using Stock Synthesis 3 (3.30.03.05, released May 11, 2017 for pre-STAR work and explorations using 3.30.04.02, released on June 2, 2017). Stock Synthesis has been extensively used as the main software and extensions have been validated and documented.

BACKGROUND

*The main body of the reviewer report shall consist of a **Background**, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.*

Yelloweye rockfish

Yelloweye rockfish (*Sebastes ruberrimus*) is a long-lived rockfish distributed from the western Gulf of Alaska to Baja California. It inhabits rocky reefs with little evidence of ontogenetic changes with depth or latitude and is regarded as sedentary, and an aggressive top predator susceptible to hook-and-line gear. The lifespan of Yelloweye rockfish is over 100 years, with a reported maximum of 154 years. Fifty per cent maturity is reached *circa* 22 years. The assessment region for Yelloweye rockfish considered in this review covers the US west coast ranging from the USA-Mexico border, through Oregon to the USA-Canada border. Although there are likely linkages with Yelloweye rockfish in British Columbia, these are assumed to be negligible. The assessment assumes a single stock but with two spatial areas (California and Oregon-Washington).

Because of its large size and fillet quality, Yelloweye has historically been highly prized by both commercial and recreational fisheries. Commercially, it has been caught by both trawl and hook-and-line. Catches increased steadily through the first half of the 20th century, with a brief peak around World War II due to increased demand. The largest removals were in the 1980s and 1990s, reaching a maximum of 552mt in 1982. In 2002, yelloweye was declared overfished. Since then, annual catches have been maintained at about 10mt from all fisheries across all areas. Yelloweye is currently caught only incidentally in commercial and recreational fisheries.

Yelloweye rockfish was last assessed in 2009, with an update in 2011. At that time, the point estimate for depletion of the spawning biomass at the start of 2011 was 21.4%, having increased from an estimated 16.1% in 2002. That assessment modelled three areas (as opposed, now, to two); estimated natural mortality (M) and recruitment steepness (h), while now both are presented to the model; modelled two sexes whereas a single sex model is now used; and did not estimate recruitment, whereas now these are estimated.

Yellowtail rockfish

Yellowtail rockfish (*Sebastes flavidus*) is a midwater, aggregating species with a lifespan of the order of 60 years. Genetic studies suggest there are two distinct stocks, separated around 40° 10'N (Cape Mendecino, CA). The Pacific Fisheries Management Council (PFMC) manages on this two-stock basis, with the northern stock subject to catch limits and the southern stock part of the "minor Shelf Rockfish" complex. Although there are likely linkages with Yellowtail rockfish in British Columbia, these are assumed to be negligible in the northern stock assessment.

Rockfish fishing history is complex, with different recreational and commercial interests on the northern and southern stocks. For both northern and southern stocks, but especially southern, there is strong recreational interest and history. The northern stock commercial fisheries developed post war using

both bottom trawl and midwater trawl. Directed midwater trawling was high during the 1980s and 1990s. Yellowtail is caught in association with other rockfish species, including canary rockfish (*Sebastes pinniger*). Concerns about other rockfish species have led to multiple restrictions on yellowtail fishing. Bottom trawling was effectively curtailed in 2000, and since 2002 closures and various other species catch limits have effectively closed all directed midwater fishing for yellowtail. Since 2011, with improvements in the status of many constraining species, opportunities for catching yellowtail have increased.

The northern stock of yellowtail rockfish was last assessed in 2005, using a three-area model implemented using ADMB. At that time, the point estimate for depletion of the spawning biomass at the start of 2005 was 55%, but variable within the three areas. Due to data revision and aggregation, the current model, implemented using SS3, is for a single area. This is consistent with genetic evidence.

The southern stock was last assessed in 2013 using the DB-SRA data-poor method. The assessment was not reviewed. No estimate of depletion was made with the method providing only an estimate of the OFL contribution to the minor Shelf Rockfish complex. This year, an age-based assessment was attempted but was unsuccessful.

REVIEW PROCESS

Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations

The Stock Assessment Review (STAR) Panel for yellowtail and Yelloweye rockfish took place at the NWFSC, Seattle, Washington, from 11th to 14th July 2017. The review was hosted by NWFSC.

Participants in the review are listed in Appendix 3. The STAR Panel comprised a PFMC SSC appointed Chair (CFiled), a NWFSC appointed reviewer (Budrick, CDFG) and two CIE reviewers (Apostolaki and Stokes). The rapporteurs (Apostolaki, Budrick and Stokes) for the STAR Panel reports were agreed on the first morning. The STAR Panel was tasked with providing separate reports for yellowtail rockfish (containing both northern and southern stocks) and Yelloweye rockfish. Notification of the meeting and dissemination of papers followed closely the schedule laid out in the CIE Statement of Work (see Appendix 2). Materials were provided in advance *via* a dedicated ftp server (see Appendix 1). Overall, administration of the review was sound.

The Terms of Reference (ToR) for the review are given in Appendix 2, Annex 2. Often, reviews including CIE experts focus on a particular phase of the stock assessment process – either the data inputs or the assessment *per se*, and often deal only with a single stock. The ToR set for the STAR 2 review (and other STAR meetings) is very wide, spanning effectively for three stocks, data quality (including collection and analysis) and the stock assessment. Given the scope of the ToR and review meeting, it was not possible to devote as much time as would be desirable to every issue area and some of the data issues perhaps did not get resolved. This is a persistent problem and was discussed at some length during the STAR. There is some frustration that STATs may not always have data available in a timely fashion, or have the right access to expertise to analyze data to derive inputs to the stock assessments. There is also some concomitant frustration with the amount of review time that might be spent on data inputs rather than

stock assessment modelling. This is a difficult issue. On the one hand, it is important that stock assessment analysts are familiar with, and understand data sources. On the other, specific knowledge and expertise to allow sensible data analysis may lie outside of STATs. Preparation of data inputs in advance in a separate process (as, e.g., done in SEDAR) might be efficient, but it is inevitable that during the stock assessment modelling, issues will arise and possible re-analysis might be warranted. Each NMFS science center deals with the process differently and there is no one, single best process. Despite the frustrations apparent during this STAR for yellowtail, but not for yelloweye rockfish, my view is that the STAR process is generally sound and, despite the problems, my experience of many STAR meetings is that it usually works well. While it is worthwhile for the regional center coordinators to compare review processes and perhaps seek to modify where necessary, I think the STAR process is sound and there is no urgency for change. The key thing is to ensure good communication and support of STATs throughout their periods of preparation.

The meeting followed the general outline of the draft agenda (Appendix 2, Annex 3), but out of necessity was fluid. Daily meetings generally started at 8:30am and continued to around 5:30-6:00pm. There was limited time within the meetings for Panel-only discussion or for real time report drafting and stock-taking of notes. Some Panels rely heavily on discussions in the evenings or over breakfast but, unfortunately, logistics prevented this happening in this case. Preparation of the two Panel reports relied on three Panel members taking extensive notes in real time, a difficult task when also attempting to pay full attention to and participate in fast moving technical discussions and to deal inter-changeably with three separate assessments with no breaks between. However, a session late on day 4 and early on day 5 helped considerably with STAR report drafting. At the close, the chair was left with three draft report sections, all discussed by the whole Panel, and the task of bringing them together to meet the scheduled deadline.

I note that the STAR was for three stocks, not two as expected. This is because the yellowtail rockfish assessment comprised separate northern and southern stocks. The southern stock assessment is completely new and the northern used SS3 for the first time. At times, I found the consideration of both stocks to be confusing, with requests being responded to for both during the STAR. The separation of rapporteuring and STAR report drafting helped, but I suggest a restriction to two stock assessments per STAR. However, the materials available pre-STAR for yellowtail were not highly detailed, raising concerns and triggering advance e-mails. The STAT responded, outlining the lateness of receipt of multiple data sets; these explain some of the omissions but not, for example, lack of detail on fishery-independent data. These are internal NWFSC issues and not within scope of review. Nevertheless, the lack of detailed advance materials did not make for an easy review and stood in stark contrast to the yelloweye pre-STAR documentation and presentation. This is especially true of descriptions about datasets but also about model fitting. Data discussions during the STAR were often very detailed, with NWFSC and State agency staff involved – as an external reviewer I found them extremely difficult and do not think I have a complete picture. The STAT for yellowtail should be congratulated on the extra materials brought to the STAR and attempts to react to pre-STAR comments.

The STAR process involves the STAR Panel working with Advisors and the STAT, not just to review data and assessments, but also to agree to final decision tables. For both yelloweye rockfish and yellowtail (north) rockfish, the Panel reached a point where agreement was reached on what would constitute base case runs and sensitivity tests, and what would go into decision tables. MCMC convergence was

checked for both stocks and draft, final decision tables were run and checked. There is no agreed stock assessment for yellowtail (south) rockfish and therefor no decision tables. Both STATs were highly responsive and all requests received clear and timely responses.

The two STATs were comprised of NWFSC staff. Panel advisors included representatives from the PFMC, the Groundfish Management team (GMT), and Groundfish Advisory Panel (GAP). Other participants included NWFSC and state agency staff as well as fishing industry. A webinar was used throughout the STAR, but I do not have a record of participation. I am not aware of any problems with notification of the meetings. All participants were able to contribute throughout the meeting. Many participants contributed usefully to discussion, and I believe that all were provided appropriate opportunity for involvement both during the Panel meeting and during extra-mural discussions.

Although it is often said, it is worth repeating that the Council and NWFSC staff involved should all be thanked for ensuring an excellent process. So too should the chair of the STAR for a light but efficient touch; the meeting was cordial throughout and met its mandate.

REVIEWER'S ROLE IN THE REVIEW ACTIVITIES

The main body of the reviewer report shall consist of a Background, ***Description of the Individual Reviewer's Role in the Review Activities****, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.*

The role of the reviewer is set out in the CIE Statement of Work, Attachment A, attached here in Appendix 2, Attachment A. Both CIE reviewers are tasked with producing an independent report to the CIE. The reviewers were additionally tasked with contributing to Panel Reports for each of yelloweye and Yellowtail rockfish.

In addition to *becom(ing) familiar with the draft stock assessments(s) and background materials*, I (Stokes) participated in all discussions for yellowtail and yelloweye rockfish, and wrote the first draft of the STAR Report for yelloweye rockfish during the meeting. The draft, along with that for Yellowtail rockfish, was considered by the Panel and left with the Chair at close on 24th June for further work. At the time of writing this (CIE) report, the Panel Reports have yet to be finalized.

SUMMARY OF FINDINGS BY STOCK

The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, ***Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.***

ToR 2 *Comment on the quality of data used in the assessments including data collection and processing.*

Yelloweye rockfish
Note that because of similar growth patterns, a single sex model is used.

Natural Mortality: There are insufficient data to inform natural mortality estimation within the model. It is common practice at the NWFSC to use estimates of natural mortality (M) based on the meta-analytic approach developed by Hamel (2015). This approach was used for the yelloweye stock assessment using the most recently available data. The STAT also explored alternatives using a tool developed by one of the STAT (Cope; (http://barefootecologist.com.au/shiny_m.html). The median of the M prior derived using the Hamel approach is 0.44 (based on the second oldest observed/read age of 23 years from the CA-OR-WA region. Discussions during the STAR considered alternatives based on 12.5 and 87.5 percentiles from the Hamel prior, use of the 99th percentile from the distribution of all observed ages (97), and the maximum observed age (154), likely from Alaska. Consideration of M as a model input was effectively considered in sensitivity testing (see below) and in formulation of states of nature for the decision tables (see below). Ageing error/bias was also considered (see below).

Steepness: There are insufficient data to inform stock-recruitment estimation within the model. A Beverton-Holt relationship is assumed and the steepness (h) parameter was fixed at 0.718, being the mean of the prior derived from the 2017 NWFSC annual meta-analysis of Tier 1 rockfish assessments. This is standard/common practice at the NWFSC. Combined with likelihood profiling and sensitivity testing to previously used h values and attempts to estimate h, this is appropriate.

Length/Age/Growth: Growth is estimated in the stock assessment for males and females combined. Length composition data and some age data are available variously from commercial and recreational fisheries. Length data are available from all three surveys - AFSC Triennial, NWFSC Trawl, and IPHC longline surveys – but age data are available only from the latter two and it is not clear that the IPHC survey is representative of the population. Overall, age and length composition data, while being the primary drivers of the assessment, are sparse. Given fishery closures since 2002, any future compositional data will need to come from surveys, most notably the NWFSC trawl. Treatment of the data as described in the pre-STAR report and during the STAR is sound. Of particular note, is a lingering issue with age reading (and hence ageing error specification in the model). Otolith reading has been carried out both by NWFSC and WDFW. NWFSC has aged NWFSC Trawl survey otoliths and Californian recreational data, with WDFW ageing all other samples. Cross-reads between WDFW and NWFSC age estimates has demonstrated potential bias in age estimation methods by one or another of these laboratories, particularly at ages 32 and above for which WDFW consistently ages fish older than NWFSC (and ADFG) by about 20 years on average. The assessment model does not use ages to estimate natural mortality and all sensitivity tests suggests status determination is little affected when alternative ageing errors are used. However, model results and projections are sensitive to the assumed value of natural mortality (see ToR 4) and refinement of maximum age readings is therefore important. As noted by the STAR, additional research efforts to confirm age estimation results using age validation methods (such as bomb radiocarbon or lead 210 validation) would also be beneficial in resolving such differences.

Maturity: The assessment uses the most recent data and analyses. Data used are from 2002 to 2015 from a variety of locations and sources. Analyses include allowance for potential false spawning events. A logistic maturity function was fit and used in the pre-STAR assessment, but this was replaced during the STAR with a slightly different ogive estimated using a smooth spline. Sensitivity testing with the

ogive used in 2011, and the new ogive (without spline) used in pre-STAR, showed the updated ogive makes very little difference to the assessment.

Surveys: Fishery independent indices are available from AFSC triennial trawl surveys conducted between 1980 and 2004 and from annual trawl surveys conducted by the NWFSC between 1999 and 2017. Neither trawl survey encounters many yelloweye rockfish. A third index is available from the IPHC fixed station longline survey incorporating additional rockfish stations from 2007 to 2016.

For many stocks assessed by the NWFSC, the triennial surveys are split between 1992 and 1995 due to changes in the timing and depth range of the survey. For yelloweye, this is not a problem as the extension of the triennial survey is to depths beyond those in which the species is found. The series is therefore treated as a single index. For both trawl surveys, only the OR-WA areas have sufficient number of positive hauls (for yelloweye) to allow analysis. For the CA area, positive hauls range from 0 to 7 per year for both surveys and are insufficient to allow analysis. From first principles, neither trawl survey is likely to be a good index – the gear is inefficient at sampling the species; the random design is arguably not ideal for the species which is highly associated with fixed biogeographical features; etc.

As described in the pre-STAR report, the trawl survey indices are estimated using the Vector Autoregressive Spatial Temporal (VAST) delta-model (Thorson et al. 2015). The method has been reviewed elsewhere and adopted by NWFSC – it is outside the scope of this review. My understanding is that the method allows a user-defined degree of geospatial smoothing and estimates both a fixed geospatial component and an annual one to allow for stocks that may be changeable annually. The method results in lower estimator error than standard techniques, but it is unclear how bias may result for different types of stock characteristic (e.g. sedentary and associated with features *cf* mobile and variable annually). I have not followed the detailed fitting statistics shown in the pre-STAR report, but accept that the application of VAST is the best available. I note that all any survey index fitting does, in effect, is fit a surface which is integrated to derive an annual estimate of biomass. With VAST being a relatively new technique, it might be more convincing to visualize how the surface looks annually and how variations in the user-definitions of knots (and presumably smoother distance, etc.) might impact the surface estimations. Figures 13 and 14 of the pre-STAR report do show the final density plots by year but are on a scale that is very difficult to interpret and apparently includes the full southerly extent of surveys when the stock in question is the northern one. Given the strong association between yelloweye and rocky features, how do the surfaces relate to known features? Do the 250 knots cover the whole survey area or just a portion (i.e., what is the scale of smoothing and how does it relate to the scale of features associated with yelloweye?).

The IPHC survey and all CPUE indices (below) are estimated using standard delta-GLM approaches with a clear specification on interpretation of statistical diagnostics provided through the SSC. The STAT thoroughly analyzed the IPHC data including the additional stations carried out for rockfish. After much filtering, the resulting statistical fit seems appropriate though it should be noted the Q-Q plot displays a marked departure from normality on the righthand side from not much above the median. While the best fit, the lognormal may still not be very good. I like, however, that only gamma and lognormal distributions are considered. It is increasingly common to see multiple distribution options offered and a choice made entirely on visual inspection of the Q-Q plots. Not all distributions are appropriate regardless of fit and the restriction to lognormal and gamma is sensible. Also, it is good to see that the

STAT has carefully considered all the available diagnostics in deciding a best fit. Overall, the IPHC fit has a high cv (*circa* 0.5).

It is hard to interpret the fishery-independent indices. The triennial survey and NWFSC trawl surveys are unlikely to provide good indices from first principles. The strong, apparent decline in the Triennial index from the late 1980s to 1995 superficially reflects the consensus view of the stock, but could be artefactual. The NWFSC shows an apparent increase, though uncertain, from 2004 until 2014, consistent with expectations after the fishery closure, but then a sudden drop. The IPHC survey is in principle better, being a hook and line survey, but is an add on in restricted areas and is very noisy. Despite the amount of careful analysis by the STAT, none of the surveys influence the assessment outcomes (see below).

CPUE: As for the IPHC survey, standard delta-GLM methods were used to analyze commercial and recreational CPUE. The STAT provided a thorough description and analysis of each data set and provided full statistical diagnostics and explanations for model choice. The results are somewhat underwhelming, with some fits (CA MRFSS Dockside, CA onboard observer, OR MRFSS dockside) seemingly acceptable and others (OR onboard observer, OR ORBS dockside, WA dockside) not. All indices are in any case flat, even through the period when the stock was heavily fished and considered to be in decline. The sensitivity tests show lack of influence in assessment results and the STAR Panel discussed whether indices might be removed from the model given the apparent lack of information they contain. It was decided to include all indices as a "place holder" for future assessments. I am not entirely comfortable with this approach. A great deal of work is expended on analyzing indices as part of the assessment process. Whether this is done by the STAT or in advance as part of a data process, it is important to make the process efficient. In my view, where data have been thoroughly analyzed and cannot be added to or improved, but have been found uninformative, it would be good practice not to burden future assessment processes. This does not mean that the data should be discarded completely, but some filtering of data to improve efficiency is warranted. This relates somewhat to the discussion on best processes to separate or meld data and assessment modelling steps.

Removals: Total removals (landings and discards) are important, especially when virgin biomass based reference points are used. The yelloweye rockfish assessment is fairly standard in that productivity parameters are assumed and a model and fitting procedure is applied to composition data from a period relatively late in the fishing history. Assumed and estimated parameters are applied in a simple accounting exercise, taking account of withdrawals (removals) over a prolonged period to estimate initial capital (virgin biomass) *circa* 1890. Any uncertainty in removals can therefore be a major source of uncertainty in estimates of depletion (as a ratio of current to virgin biomass). Estimating removals from three states, for commercial and recreational fisheries, is a challenge. A considerable amount of the STAT effort and reporting is involved with the estimation of removals. The effort involves the STAT (and previous STATs), multiple science centers, state agencies, etc. As a general point, this is perhaps the most important step potentially to be considered outside of the main assessment process, and it is clear that there is an ongoing and strenuous effort by all involved for all species. Reporting on removals as part of the STAR process is important, but the details of recreational and commercial catch monitoring schemes are well outside the scope of most STAR Panels. Taken at face value, the pre-STAR report and discussions during the STAR meeting suggest the removals data (and composition data) by fleet are the best available for stock assessment. Even during review, discussions with all present led to some

relatively minor requests for data tidying that were incorporated into the final assessment. None made any difference of note to model fitting or outcomes.

Overall, I am content that the quality of data used for the stock assessment is sufficient to allow rigorous modelling and development of management advice.

Yellowtail rockfish

Northern and Southern stocks

Natural Mortality: It is common practice at the NWFSC to use estimates of natural mortality (M) based on the meta-analytic approach developed by Hamel (2015). This approach was used for the pre-STAR yellowtail stock assessment explorations using the most recently available age data to provide medians of M priors for female and male M for both northern and southern stocks. 99th percentiles for males and females, for both stocks, from early time-series (unclear) were used to start explorations. The final models used the female 99th percentile to fix the median for the northern stock and allowed a male offset during M estimation. Southern stock assessments "borrow" estimated values from the northern assessment.

I note that section 2.3 and 1.2 of the pre-STAR report give respectively, female and male 99th and 95th percentiles of age for both stocks. However, the age values given are the same in both sections. I cannot find an age distribution elsewhere and this is unclear. I suggest it is clarified.

I note also that in the pre-STAR northern model fits (see e.g. slide 15 of the presentation on the yellowtail (north) model), the likelihood profiles suggest age composition data consistent with higher M than the Hamel-derived median for female M, but that the index data (though not well fit) push the M estimate towards the median of the prior. This influence through the index data is not intuitively reasonable and masks the difference between the input median and information in the age compositions. The likelihood profiles also suggest strong age composition support for a male M offset smaller than that implied by maximum ages. More interesting is that the maximum age observations suggest higher and more sex divergent M in the southern stock.

Overall, and knowing in advance that M estimation will be the major axis of uncertainty, these are concerns.

Steepness: There are insufficient data to inform stock-recruitment estimation within the model. A Beverton-Holt relationship is assumed and the steepness (h) parameter was fixed at 0.718, being the mean of the prior derived from the 2017 NWFSC annual meta-analysis of Tier 1 rockfish assessments. This is standard/common practice at the NWFSC. Combined with likelihood profiling and sensitivity testing to previously used h values and attempts to estimate h, this is appropriate.

Maturity: The assessment uses the most recent data from a variety of sources. Sample sizes are small and insufficient to allow area specific estimation.

Length/Age/Growth: Growth is estimated in both the northern and southern stock assessments females, with offsets for males. Length composition data and some age data are available variously from commercial and recreational fisheries.

For the northern stock, length data are available from two trawl surveys (AFSC Triennial, NWFSC Trawl) but age data are available only for the Triennial survey. Neither bottom trawl survey is regarded as a good sampling system for yellowtail which is generally found in midwater. Conditional age-at-length compositions are available for the NWFSC survey. Overall, age and length composition data, while being the primary drivers of the assessment, are sparse, especially from commercial fisheries which have been impacted by closures and constraints on other species. For recreational fisheries, there is the added complication of low numbers of older fish perhaps being unselected by the gear. Analysis and treatment of the data as described in the pre-STAR report is poor and STAR discussions were difficult to follow. I note that issues with data coding (for the model) and with data cleaning (e.g., WA recreational fishery units) were dealt with during the STAR and incorporated into final model runs. I do not feel able to comment on the data as such, but am content that the STAR provided a thorough examination.

For the southern stock, there are length data for recreational and commercial fisheries since 1980 but age data are only available from the commercial fishery. There are length data available from a hook and line survey since 2004, but with age data only for 2004; the survey only covers the southern part of the stock distribution. The paucity of age and length data means the assessment relies on growth estimates (and other life history information) from the north. The starting point for an age-based assessment is clearly fragile. If the southern stock is to be assessed separately using an age-based method, ageing needs to be carried out for the hook and line survey, and that all sources are sampled in the future. The survey perhaps offers the best opportunity both for ageing and for enhancing life history studies to understand southern *cf* northern values.

Surveys: For the northern stock, fishery independent indices are available from AFSC triennial trawl surveys conducted between 1980 and 2004 and from annual trawl surveys conducted by the NWFSC between 1999 and 2017. Both surveys use bottom gear and neither is ideal for sampling yellowtail in midwater. As noted by the STAR, alternative survey methods (acoustics, midwater trawls) could provide better indices of stock abundance. While I agree that such methods might be considered, it is beyond the scope of this review to recommend them – that is dependent on the utility of such surveys for a broad set of species and relevant cost-benefit considerations.

As described in the pre-STAR report, the trawl survey indices are estimated using the Vector Autoregressive Spatial Temporal (VAST) delta-model (Thorson et al. 2015). Please see comments for yelloweye rockfish.

It is not possible to comment on the southern stock hook and line survey as no information was provided.

CPUE: For the northern stock, the STAT used standard approaches for most commercial CPUE and for recreational CPUE, but used a VAST approach for an index for yellowtail bycaught in the hake fishery. The STAR revealed problems with all indices. For the commercial indices, it became apparent that not all states use the same identifier codes and that filtering as done could be problematic. A (possible) issue was

identified in that the CPUE analyses fit to catch/tow and a request was made during the STAR to include time in the dependent variable. I am not worried by this as duration is included in the pre-STAR model. In any case, the resulting indices are little affected. A request was also made for fuller documentation on the VAST approach to the hake bycatch CPUE. In discussion, the STAT and STAR agreed that none of these indices should be used in the final northern yellowtail model at this time, but that all warrant further investigation.

Removals: See general comments for yelloweye rockfish. The pre-STAR report did not include any detail on removals for either stock. The STAR did not have any discussions on specific issues. I have no comment.

Overall, I consider there to be sufficient stock specific life history information, and compositional data to allow a stock assessment for northern yellowtail. It is unfortunate that there is no good index of abundance. For the southern stock, the lack of stock specific life history information and limited age composition data suggest an age-based stock assessment is not yet possible. The lack of a good abundance index is a further complication that would affect even simpler modelling approaches.

ToR 3 *Evaluate and comment on analytic methodologies.*

Yelloweye rockfish

The assessment of yelloweye rockfish was implemented using Stock Synthesis 3 (3.30.04.02; released on June 2, 2017). Stock Synthesis has been extensively used and the main software and extensions have been validated and documented.

The assessment model assumes a single US west coast-wide stock of Yelloweye rockfish but with two areas (CA and OR-WA) and seven fleets. Catch data are split by fleet but not by sex. Recruitment for the stock is distributed annually to the two areas assuming a fixed proportion. M, h, maturity, fecundity, and growth are externally supplied.

The catch history starts in 1890. Catches are provided by fleet, including discards (which are regarded as small). The model estimates separate selectivity for each fleet. While allowed to be domed, commercial and recreational fleets are all estimated as asymptotic. Selectivity blocking was considered, but not used due to lack of data in the period for which it was contemplated (post 2002). Two trawl and one longline survey are fit, as are six fishery-dependent CPUE. None of the indices are influential – two trawl surveys are not expected to be informative and the fishery-independent indices are uninformative. Age and length data and conditional age-at-length compositions are the most influential data sources.

The model was explored using SS3 and the pre-STAR report is excellent. The report and the presentation to the STAR include extensive explanation about model changes since 2011 (with clear rationales for each), fit diagnostics, likelihood profiling, and sensitivity/influence summary plots. The pre-STAR base model was also run to full MCMC, with jittering and standard tests. The sensitivity/influence summaries show relative errors in management-related model outputs for all indices, lengths, and ages, as well as for ageing error, M, h, maturity and fecundity, catch history, and miscellaneous items such as selectivity

type and weighting methods. The sensitivity/influence plots are excellent and I recommend they be standardised and used in all assessments where possible. Of course, their presentation pre-STAR requires comprehensive and organised work. The STAT should be congratulated on what it achieved pre-STAR.

During the STAR, the few exploratory model runs were to MPD only. Tuning of model runs followed standard (Francis) procedures with multiple passes to refine index and composition weights. A full MCMC has been conducted on the post-STAR candidate base case and there do not appear to be any problems with convergence. None would be expected given externally input M, h and growth.

The model is relatively simple and has been extensively explored. It is a well-rationalized simplification from the previous benchmark assessment model and the assessment provides a robust basis for decision making with the decision tables using a reasoned range of M to capture model misspecification – a greater source of uncertainty than is captured within any single model run.

Yellowtail rockfish

The assessments of northern and southern yellowtail rockfish stocks were implemented using Stock Synthesis 3 (3.30.03.05, released May 11, 2017 for pre-STAR work and explorations using 3.30.04.02, released on June 2, 2017). Stock Synthesis has been extensively used and the main software and extensions have been validated and documented.

The pre-STAR reports for both stocks are not detailed, particularly with respect to data preparation and justification of choices made. This makes evaluation difficult.

The northern stock assessment assumes a single area, sex-specific life histories, and four fleets. Length and age compositions are available with conditional age-at-length available only for the NWFSC survey. None of the indices of abundance are well fit and there are no clear signals; they do, however, influence M estimation, generally "pulling" female M lower than composition data. This is clear both from pre-STAR likelihood profiles and from one request (to run the model without indices) made during the STAR. The main fishery-dependent CPUE index is for a combined commercial fleet, including bottom and midwater trawls; however, the fitted GLM includes multiple variables/factors which could alias for gear type and this may not be a problem. The surveys both use bottom gear, but yellowtail is a midwater species. In general, given the lack of a clear index of abundance, the STAT has pre-STAR and during STAR, explored sufficiently to demonstrate the utility of an age-based assessment relying predominantly on age-at-length from one survey and input life history parameters.

The identification of M as the main axis of uncertainty followed consideration during the STAR of within model error and of likelihood profiles on R0 and M. The values chosen for the decision table are the final run estimated M and the 12.5 and 87.5 percentiles of the prior on M re-centered on the estimate. The values are 0.122, 0.159, 0.249.

During the STAR, the few exploratory model runs were to MPD only. Tuning of model runs followed standard (Francis) procedures with multiple passes to refine index and composition weights. A full

MCMC has been conducted on the post-STAR candidate base case and there do not appear to be any problems with convergence. None would be expected given externally input M, h and growth.

The model is relatively simple and has been sufficiently explored. It is a simplification (one *cf* three areas) from the previous benchmark assessment model and the assessment provides a robust basis for decision making with the decision tables using a reasoned range of M to capture model misspecification – a greater source of uncertainty than is captured within any single model run.

For the southern model, the assessment would best be viewed as exploratory, using standard analytical and modelling approaches as applied to other stocks to test feasibility and highlight data deficiencies. The limited, available age data suggest a higher natural mortality than for the northern stock. Based on observed maximum ages alone, this may be feasible, but inferences need to be circumspect. In the absence of a good abundance index and with few age data, especially in recent years, the model is highly unstable and does not provide a basis for management advice. I can see no issue with the choice of analytical approaches to data preparation or assessment, except that they have been applied to poor data, and in the case of data preparation, have suffered from timeliness of data provision and a possible lack of familiarity with some of the data. The STAT has tried a variety of model configurations, has considered diagnostics, and has explored sensitivities, but the data simply do not support an age-based assessment at this time. As noted above, it is problematic that there is no good index of abundance as would be required for simpler assessment approaches.

ToR 4 *Evaluate model assumptions, estimates, and major sources of uncertainty and provide constructive suggestions for improvements if technical deficiencies or additional major sources of uncertainty are identified.*

Yelloweye rockfish

The model assumes a single, distinct US west coast stock with no linkage to yelloweye rockfish beyond US waters. This may be a simplification but is not likely a major problem for assessment and management purposes unless linkages are such to render the stock-recruit assumptions invalid. I would give this a low priority, but it would be useful to see a comment in the final report.

The single-stock and two-area model assumes a fixed, annual redistribution of recruits (60% to OR-WA and 40% to CA). The reason for this split is not clear in the pre-STAR report but emerged as one issue for consideration during the STAR. A request was made by the Panel to run the model assuming an 80:20 split as a sensitivity test. The STAT went further and helpfully provided likelihood profiles on the ratio which clearly demonstrated a model preference for the 60:40 ratio, strongly driven by the length data from California and Oregon. My view is that the ratio should not be considered as biologically meaningful but is more a book-keeping device within the model. I suggest that further exploration should **not** be a priority as there are no data to inform other than a single, annual ratio. Also, the single stock sensitivity test showed little influence on management-related model outputs and the chosen states of nature already capture wide model specification uncertainty.

The model currently uses two areas because of difficulties is assigning catches and composition data between Oregon and Washington. It is not clear that work to allow further disaggregation will necessarily improve the model as a tool for informing management decision-making. Splitting the assessment to accommodate more areas might be beneficial if management at a finer scale is desirable, but as a single stock with no detailed fishery-independent information at finer scales, the assessment would in any case be driven largely by assumptions. This is not to say that improving historical catch estimates is not warranted. All assessments benefit from better data whether model outputs are affected or not, if only because the model and emergent advice becomes more credible. To the extent possible, therefore, I would encourage constant improvement in line with recommendations elsewhere on catch reconstruction and as clearly taken seriously in the region.

Fundamentally, assessment models are just ways of weighting different data sources to extract and balance information in a structured framework – to provide management-related outputs. Use of the outputs is best achieved when there are clear consistencies and where the raw data relate directly to the outputs. Fishery independent survey indices are therefore especially important in creating credibility because they have a simple interpretation. The more outputs are driven by technical decisions on e.g. weighting, on assumptions e.g. about productivity, or on hard-to-discern patterns in composition data, the more tenuous the results. In this case, there are no credible indices and the assessment is driven somewhat by weighting choice (but tested for), by composition data that generally tell a consistent story and display no bad residual patterns, but especially by the assumptions about productivity – whether through h or M. The choice of M and the decision table flanks are based on ageing.

Data are limited and the lack of a reliable abundance index leaves the assessment (and management) dependent on composition data and assumptions. It is unlikely that any of the CPUE or trawl survey indices will prove useful in future and it is unclear if the IPHC additional stations may be informative. This is unfortunate but real. The composition data are limited and, as ever, do not span early fishery history. Nevertheless, they appear reasonably informative and jointly, though not by individual data set, are the only data that influence management-related model outputs. Further refinement and continued collection is imperative. Most importantly, however, given the primary sensitivity is to the choice of M, further work to understand ageing and maximum age is an obvious research priority. As noted above, application of age validation methods such as bomb radiocarbon or lead 210 would be beneficial.

Yellowtail rockfish

While life history data for the northern stock seems reasonable, further data collection from the southern stock (and Canada) is warranted to allow better functional stock separation and to feed into stock assessments. Studies could consider not just regional variation but also age- and sex-specific natural mortality.

<u>Northern stock only (as no southern stock assessment concluded)</u>

The model assumes a single, northern US west coast stock with no linkage to yellowtail rockfish in Canadian waters. This is clearly a simplification and there is a commercial fishery off British Columbia. As noted above, this is unlikely to be a major problem for assessment and management purposes unless

linkages are such to render the stock-recruit assumptions invalid. Investigation of the issue would require access to US and Canadian fishery and biological data. I would give this a low priority but it would be useful to see a comment in the final report.

The lack of a credible abundance index is a weakness in the assessment. With little likelihood of a credible index being derived from the existing bottom trawl surveys, and no immediate prospect of acoustic or midwater surveys (see also ToR 6), work on improving CPUE indices should be undertaken before any assessment update. The indices are currently poorly fit but, because of data paucity, are influential. It is perplexing to hear during a STAR that the data are not in easy order for analysis and that analysts are not completely familiar with the data collection schemes and fisheries. CPUE analysis is not just a statistical exercise. I recommend that future CPUE analyses involve appropriate State and fishery personnel to ensure full understanding of data - even if this means a separate, data preparation step in the STAR process.

The model assumes natural mortality is constant for males and females. Given the use of spawning output as opposed to spawning biomass, an exploration of age-related natural mortality and the implications of using it might be warranted. This relates also to selectivity estimation and the choice of appropriate forms for surveys and CPUE. With indices generally being poorly fit, it is possible that the model could be enhanced by a better understanding of the mortality schedule.

ToR 5 *Determine whether the science reviewed is considered to be the best scientific information available.*

Yelloweye rockfish

In my opinion, the STAT has comprehensively reviewed the available information on Yelloweye rockfish and has conducted thorough analyses to provide estimates of management-related quantities. Uncertainties in inputs and estimates of interest have been exceptionally well explored and presented, and I am confident that the resulting assessments and decision tables represent the best scientific information available.

Yellowtail rockfish

The STAT has reviewed the available information on yellowtail rockfish and has conducted analyses to provide estimates of management-related quantities for the northern, but not the southern stock. The STAT should be congratulated on attempting the assessment of the southern stock but the data simply do not support an age-based stock assessment and the borrowing of life history information from the northern stock is counter-intuitive to the noted differences. Uncertainties in inputs and estimates of interest have been explored and presented for both southern and northern stock assessments, and I am confident that the resulting northern stock assessment and decision tables represent the best scientific information available for management purposes. I do not consider the southern stock assessment to be usable.

ToR 6 *When possible, provide specific suggestions for future improvement in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.*

For all stocks, the most pressing need is for a credible index of abundance, ideally fishery-independent. The STATs made good attempts to use standard delta-GLM and VAST analyses of CPUE and survey data. In all cases, however, there were problems. In all cases, the surveys are generally inappropriate or unlikely to provide valid abundance indices regardless of the quality of analysis. I am reluctant to recommend specific surveys for specific stocks (e.g., acoustic for the northern yellowtail stock) because survey needs vary by stock and planning needs to take account of multiple stock types and cost-benefits. I suggest that the feasibility and cost-benefits of alternative survey types (midwater, line, and acoustic) be estimated for stocks not well served by the bottom trawl surveys.

Equally, it is clear that lack of good age composition data is a limitation for many stock assessments. For southern yellowtail, the lack of almost any recent age data made an age-based stock assessment untenable. For the other stocks considered here the composition data were sufficient but could be strengthened. Again, I am reluctant to make recommendations because all stocks and cost-benefits need to be considered. I suggest that benchmark assessment scheduling and planning should take account of available data and put in place as necessary ageing work to enable stock assessment. There is little value in attempting age-based stock assessment (as for southern yellowtail) unless that preparatory work is done in a timely fashion.

Regardless of final success or otherwise, I consider all assessments to have used appropriate data analytical methods (delta-GLM, VAST, etc., generally following SSC guidelines) and modelling techniques (weighting, profiling, considering residuals, convergence testing, etc.) and have no specific technical suggestions to make.

See also ToR 4 for specific issues related to each of yelloweye and yellowtail rockfish.

CONCLUSIONS AND RECOMMENDATIONS

The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and ***Conclusions and Recommendations in accordance with the ToRs****.*

Specific recommendations and suggestions are highlighted in **bold, red** in the preceding sections. I distinguish between recommendations as necessary activities and suggestions as desirable ones, recognizing that research planning and prioritization requires consideration of multiple factors and applies to many stocks, fisheries and other factors.

APPENDIX 1

BIBLIOGRAPHY

Prior to the Workshop, extensive materials were provided *via* a dedicated, anonymous ftp server (ftp://ftp.pcouncil.org/pub/GF_STAR2_2017_Ytail_Yeye/). The materials were extensive and relevant to all terms of reference in varying degrees, consisting of general and review-specific background materials, and draft assessments for yelloweye rockfish and yellowtail rockfish (both northern and southern stocks). For each stock, draft assessment reports plus all SS3 files and plots were provided.

During the workshop multiple presentations were given, and additional materials were provided on request, including further background documents and presentations as well as responses to Panel requests. All files were made available using the dedicated server which was accessed using guest Wi-Fi throughout the meeting. The access was generally adequate. Directory listings are not provided here as the server is anonymous (and therefore publicly available).

APPENDIX 2

Attachment A: Statement of Work for Dr. Kevin Stokes

Statement of Work

External Independent Peer Review by the Center for Independent Experts

Stock Assessment Review (STAR) Panel 2

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information on the CIE program may be obtained from www.ciereviews.org.

Project Description

The National Marine Fisheries Service and the Pacific Fishery Management Council will hold stock assessment review (STAR) panels in 2017 to evaluate and review benchmark assessments of Pacific coast groundfish stocks. The goals and objectives of the groundfish STAR process are to:

1) ensure that stock assessments represent the best available scientific information and facilitate the use of this information by the Council to adopt OFLs, ABCs, ACLs, (HGs), and ACTs;

2) meet the mandates of the Magnuson-Stevens Fisheries Conservation and Management Act (MSA) and other legal requirements;
3) follow a detailed calendar and fulfill explicit responsibilities for all participants to produce required reports and outcomes;
4) provide an independent external review of stock assessments;
5) increase understanding and acceptance of stock assessments and peer reviews by all members of the Council family;
6) identify research needed to improve assessments, reviews, and fishery management in the future; and
7) use assessment and review resources effectively and efficiently.

Benchmark stock assessments will be conducted and reviewed for yelloweye and yellowtail rockfishes. Yelloweye rockfish was assessed as a benchmark assessment in 2009 and fully updated in 2011. In 2015, a catch-only projection update was conducted to monitor this rebuilding stock and provide scientific-based advice for management. Yelloweye rockfish remains a highly constraining species for coastwide nearshore commercial and recreation fisheries, and this assessment will incorporate more recent information regarding rockfish productivity. Yelloweye rockfish has been managed under a rebuilding plan for over a decade and is not expected to be rebuilt for several more decades. However yelloweye rockfish was identified as a strong candidate for assessment during the Pacific coast groundfish regional stock assessment prioritization process, which was based on the national stock assessment prioritization framework (http://www.st.nmfs.noaa.gov/Assets/stock/documents/PrioritizingFishStockAssessments_FinalWeb.pdf.

Yellowtail rockfish was historically an important target species for mid-water and bottom-trawl fisheries. Opportunities to target yellowtail rockfish in those fisheries were greatly reduced when widow rockfish rebuilding began 15 years ago and continued early in the trawl rationalization program. Now that both widow rockfish and canary rockfish have completed rebuilding, a mid-water fishery targeting these species is expected to grow in the near future. The northern portion of the stock was last assessed as part of the data-moderate assessments in 2013, which showed the stock in the northern part of the coast to be increasing, and above the target biomass. However, the yellowtail rockfish assessment will have considerable new information relative to the last benchmark (2001) and data-moderate assessments. Benchmark assessments also involve more data types and complex modeling, and therefore, if supported by the available data, provide more complete and less uncertain estimate of the stock biomass and relative status.

Assessments for these two stocks will provide the basis for the management of the groundfish fisheries off the West Coast of the U.S. including providing scientific basis for setting OFLs and ABCs as mandated by the Magnuson-Stevens Act. The technical review will take place during a

formal, public, multiple-day meeting of fishery stock assessment experts. Participation of external, independent reviewer is an essential part of the review process. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers
NMFS requires two CIE reviewers to participate in the stock assessment review panel. One CIE reviewer shall conduct an impartial and independent peer review of the two assessments described above and in accordance with the SoW and ToRs herein. Additionally, a second "consistent" CIE reviewer will participate in all STAR panels held in 2017 and the SOW and ToRs for the "consistent" CIE reviewer are included in a separate SoW (See **Attachment A**).

Both CIE reviewers shall be active and engaged participants throughout panel discussions and able to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers shall have excellent communication skills in addition to working knowledge and recent experience in fish population dynamics, with experience in the integrated analysis modeling approach, using age-and size-structured models, use of Markov Chain Monte Carlo (MCMC) to develop confidence intervals, and use of Generalized Linear Models in stock assessment models.

Statement of Tasks
Each CIE reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: At least two weeks before the peer review, the contractor will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to the STAR Panel meeting include:

- The current draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation
- Additional supporting documents as available.
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein.

Contract Deliverables - Independent CIE Peer Review Reports: The CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

Other Tasks – Contribution to Summary Report: The CIE reviewers may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. The CIE reviewers are not required to reach a consensus, and should provide a brief summary of each reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Timeline for CIE Reviewers

The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided in advance of the peer review.
2) Participate during the STAR Panel 2 review meeting **in Seattle, WA, during the dates of July 10-14, 2017,** as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than **July 28, 2017**, each CIE reviewer shall submit their draft independent peer review report to the contractor. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**.

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the

Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
For the **STAR panel 2** review, the CIE reviewers shall conduct an independent peer review during the panel review meeting scheduled in **Seattle, Washington during the dates of July 10-14, 2017.**

Period of Performance
The period of performance shall be from the time the award through August 30, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables
The contractor shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

June 2, 2017	Contractor selects and confirms reviewers
June 26, 2017	Contractor provides pre-review documents to the reviewers
July 10-14, 2017	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 28, 2017	Contractor receives draft reports
August 14, 2017	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content in **Annex 1**; (2) The reports shall address each ToR as specified **Annex 2**; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $8,200.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contacts
Stacey Miller, NMFS Project Contact
National Marine Fisheries Service,
2032 SE OSU Drive
Newport, OR 97365
Stacey.Miller@noaa.gov
Phone: 541-867-0535

Jim Hastie
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112
Jim.Hastie@noaa.gov
Phone: 206-860-3412

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Stock Assessment Review (STAR) Panel 2

1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.
2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.
3. Evaluate model assumptions, estimates, and major sources of uncertainty.
4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Annex 3: Tentative Agenda

TBD

Stock Assessment Review (STAR) Panel 2

Pacific ocean perch and Yellowtail rockfishes
Seattle, Washington

NWFSC
2725 Montlake Blvd, NE
Seattle, WA 98112
July 10-14, 2017

APPENDIX 3
PERTINENT INFORMATION FROM THE REVIEW

1) Participants List

Reviewers:
Panayiota Apostolaki, Center for Independent Experts
John Budrick, California Department of Fish and Game
John Field (Chair), Scientific and Statistical Committee (SSC) representative
Kevin Stokes, Center for Independent Experts

Advisors:
John DeVore, Pacific Fishery Management Council (PFMC) representative
Heather Reed, Groundfish Management Team (GMT) representative
Dan Waldeck, Groundfish Advisory Subpanel (GAP) representative

STAT Members present:
Jason Cope, Northwest Fisheries Science Center
Vladlena Gertseva, Northwest Fisheries Science Center
Andi Stephens, Northwest Fisheries Science Center
Ian Taylor, Northwest Fisheries Science Center

Others:
Andrew Claiborne, Washington Department of Fish and Wildlife
Owen Hamel, Northwest Fisheries Science Center
Jim Hastie, Northwest Fisheries Science Center
Chantel Wetzel, Northwest Fisheries Science Center
Theresa Tsou, Washington Department of Fish and Wildlife
Stacey Miller, Northwest Fisheries Science Center

2) Final Proposed Agenda

Proposed Agenda
Stock Assessment Review (STAR) Panel
for Yellowtail and Yelloweye Rockfish

NOAA Fisheries, Northwest Fisheries Science Center
2725 Montlake Blvd. East
Seattle, WA 98112

July 10-14th, 2017

This meeting is open to the public and public comments from attendees will be accepted at the discretion of the meeting chair. Agenda times are approximate and are subject to change.

Monday, July 10

8:30 a.m.	Welcome and Introductions
9:00 a.m.	Review the Draft Agenda and Discuss Meeting Format (Chair) - Review the Terms of Reference (TOR) for assessments, Accepted Practices Guidelines, and STAR panel responsibilities - Assign reporting duties - Agree on time and method for accepting public comments
9:30 a.m.	Presentation of the Yellowtail rockfish Assessment - Overview of data and modeling
12:30 p.m.	Lunch (On Your Own)
2:00 p.m.	Q&A session with Yellowtail rockfish Stock Assessment Team (STAT) STAR Panel discussion - Panel develops written request for additional model runs / analyses
4:00 p.m.	Begin Presentation of the Yelloweye Rockfish Assessment (if time allows) - Overview of data and modeling
5:30 p.m.	Adjourn for day

Tuesday, July 11

8:30 a.m.	Continue and Complete Presentation of the Yelloweye Rockfish Assessment - Overview of data and modeling
10:00 a.m.	Q&A Session with the Yelloweye Rockfish STAT Panel Discussion - Panel develops written request for additional model runs / analyses
12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Begin Drafting the STAR Panel Report
4:00 p.m.	Begin Presentation of the First Set of Requested Model Runs by the Yellowtail rockfish STAT
5:30 p.m.	Adjourn for day

Wednesday, July 12

8:30 a.m.	Continue Presentation of the First Set of Requested Model Runs by the Yellowtail rockfish STAT - Q&A session with the Yellowtail rockfish STAT & Panel discussion - Panel develops request for second round of model runs / analyses for the Yellowtail rockfish STAT
10:30 a.m.	Begin Presentation of the First Set of Model Runs by the Yelloweye Rockfish STAT
12:00 p.m.	Lunch
1:30 p.m.	Continue Presentation of the First Set of Model Runs by the Yelloweye STAT - Q&A session with the Yelloweye Rockfish STAT & panel discussion - Panel develops request for second round of model runs / analyses for the Yelloweye Rockfish STAT.
3:30 p.m.	Continue Drafting STAR Panel Report
5:30 p.m.	Adjourn for day

Thursday, July 13

8:30 a.m.	Presentation of the Second Set of Model Runs by the Yellowtail rockfish STAT - Q&A session with the Yellowtail rockfish STAT & panel discussion - Agreement of the preferred model and model runs for the decision table - Panel continues drafting the STAR report.
12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Presentation of the Second Set of Model Runs by the Yelloweye Rockfish STAT - Q&A session with the Yelloweye Rockfish STAT & panel discussion - Agreement of the preferred model and model runs for the decision table - Panel continues drafting the STAR report.
5:00 p.m.	Continue Panel Discussion or Drafting of the STAR Panel Report
5:30 p.m.	Adjourn for day

Friday, July 14

8:30 a.m.	Consideration of Remaining Issues - Review decision tables for assessments
10:00 a.m.	Panel Report Drafting Session
12:00 p.m.	Lunch (on your own)
1:30 p.m.	Review First Draft of the STAR Panel Report
4:00 p.m.	Panel Agrees to Process for Completing the Final STAR Report for Council's September Meeting Briefing Book (Requested by August 14th)
5:30 p.m.	Review Panel Adjourns